

A Tangible Mixed Reality Interface for the AMI Automated Meeting Assistant

Jochen Ehnes

Centre for Speech Technology Research, University of Edinburgh,
10 Crichton Street, Edinburgh, UK
`jehnes@inf.ed.ac.uk`

Abstract. In this paper we describe our approach to support ongoing meetings with an automated meeting assistant. We propose an alternative user interface for the AMIDA Content Linking Device. In order for the system to be less distractive and more collaborative than the original laptop screen based one, we developed a system that projects documents onto the table tops right in front of the meeting participants. This way they appear as if they were printed on paper, lying in front of the participants. We describe our setup as well as the user interface we built to handle and share these documents.

Keywords: meeting assistants, meeting processing, mixed reality (MR), projected user interface, tangible user interface (TUI).

1 Introduction

Started in January 2004, the European AMI (Augmented Multi-party Interaction) integrated project, as well as its follow on project AMIDA (Augmented Multi-party Interaction with Distance Access), has been about building systems to enhance the way meetings are run and documented. AMI research revolves around instrumented meeting rooms which enable the collection, annotation, structuring, and browsing of multimodal meeting recordings. So far, meeting participants can use AMIs JFerret browser [1] to go through previous meetings and get themselves up to date. While the possibility to look through recordings of previous meetings and being able to search for important sections by keywords is a very helpful tool, it requires direct action by the user. Furthermore, as the users have to interact with the system on a personal computer, which draws their attention to it and distracts from a conversation, the browser is more useful in preparation for a meeting than during the meeting itself.

The aim of this work is to support ongoing meetings with an automated meeting assistant based on the AMIDA Content Linking Device [2]. The basic idea is that the assistant monitors the ongoing meeting. As it gets the transcript from the speech recognition system, it tries to determine the topic of the conversation, searches for relevant documents from previous meetings and displays them. While the results of this content linking process were displayed on a laptop in the original setup, we then aimed at making the system less distracting and more supportive

of collaboration. We built a system that projects documents on table tops in front of meeting participants, resembling pieces of paper. By presenting information of relevance to the current discussion in this way instead of personal computers at each participants place, we expect the ongoing meeting to be disrupted less as the participants do not have to interact with their computers. Also, a projection on the tabletop supports the communication within the group as all participants can see the projected documents and move them around easily, just as traditional documents on paper. As a result, the participants can discuss these documents in a more natural way. In order to provide enough space for several users, we use multiple projection systems, one for each user.

Recently back-projection systems with camera based multi touch screens (such as Microsofts surface computers) have become quite popular. However, they require significant volume below the surface, which would become a disadvantage for large display areas, as meeting participants would not be able to sit at the table comfortably. Furthermore we were aiming for a technology that can provide the additional functionality without requiring big changes to the existing meeting environment, as we already had a lot of equipment installed in our instrumented meeting room. For those reasons we decided to go for a system projecting from above onto the existing tables.

In order to make the interaction with the system as direct as possible, we aimed to make the projected objects graspable. Therefore we track physical objects carrying fiducial markers that serve as grabbing devices. In order to move around the projected objects, a physical document grabbing device is associated with them. As long as this connection exists, the virtual document follows the grabber. And it does so not only on one users space, but across multiple projection units. Also, this connection continues after long pauses of visibility, even between different runs of the projection system. This way important documents can be kept in reach by keeping them connected to grabbers. In a similar way keyboards can be connected to documents.

The meeting assistant brings up new documents in regular intervals, often several documents at the same time. Consequently it is not enough to make the documents appear at a default location (e.g. the center of the tabletop). We implemented a way to spread them out automatically. Of course the space on the table is not unlimited, so we scale down documents and rearrange them after a while before they get removed eventually.

In addition to augmenting the table, the system incorporates the whiteboard as well. However, as earlier meeting recordings showed, people usually have an inhibition to get up and go to the whiteboard to write something onto it. In fact, a special task had to be designed to make them make use of it. In order to reduce this inhibition, we developed an alternative way to enable participants to interact with content on the whiteboard directly from their place and move content between their space and the shared whiteboard space easily. We considered the whiteboard as a form of shared space. By making this shared space available at each projection unit in form of a second layer on top of the normal display area, the content of the whiteboard can be manipulated from every participants place.

In the following sections we will describe the concepts and details of the interaction with our system, as well as an overview of its implementation. We will also present some results from our prototype setup for two users spaces and a whiteboard.

2 Previous and Related Work

This work builds on the AMIDA Content Linking Device [2]. The system consists of a Document Bank Creator (DBC) that gathers documents that are of potential interest for an upcoming meeting, a Document Indexer (DI) that creates an index over the document bank created by the DBC, a Query Aggregator and a User Interface, all connected via a central Hub. During a meeting, the Query Aggregator performs document searches at regular time intervals using words and terms from the automatic speech recognition system. It produces a list of document names, ordered by relevance, based on the search results, as well as on a persistence model, assuming that documents that come up during several searches are likely to be more relevant than others that do not. The User interface finally displays the results to the user(s). For this work we replaced the flash based user interface with a mixed reality interface, projecting virtual documents onto the tabletop in front of the meeting participants.

The idea to use the table top as an interface to computers is not new. The first system of that kind known to the authors was DigitalDesk [3,4]. Its main intention was to bring together electronic- and paper-documents.

In [5,6] a similar setup consisting of video projector and camera, I/O-Bulb as the authors call it, mounted above the table was used to create applications that are manipulated using physical objects. Applications include the simulation of holography setups using physical placeholders for optical elements such as lasers, mirrors etc. or the simulation of fluids flowing around objects standing on the tabletop. An obvious advantage of this kind of user interface is their collaborative nature as several users can manipulate different physical objects on the tabletop at the same time, instead of being restricted to a single mouse in a conventional computer setup.

While being able to see what everybody else sees is a very important factor for collaboration, it sometimes is necessary to have some private space as well in order to work out an idea before presenting it to the whole group. In [7] the authors presented a system that supports the discussion of virtual prototypes by a group of designers/engineers sitting around a projection table. However, the difference to other 3D viewers or the applications running on the I/O-Bulbs was that the content does not occupy the whole screen space. Instead the virtual prototype would be visible on a virtual piece of paper. Apparently conventional plots of CAD drawings were still used frequently during design review meetings, as it was so easy to sketch on them with a pencil to point out flaws or suggest improvements. Furthermore, one could just work out an idea on a piece of paper before presenting it to other meeting participants. In order to make computer systems more usable for such applications, the above mentioned prototype was

developed. As the 3D models were displayed on the virtual pieces of paper, they were visible to everyone and they could be moved around using a tracked puck, so that they could be brought closer and rotated to a single person to allow for a more personal use. By grabbing two points, one with a puck and one with a pen, the virtual paper could be scaled, similarly to the two finger pinch scaling gesture known from the iPhone. Using tracked pens, participants could draw lines on the objects to annotate them. Furthermore, the system allowed to connect each piece of paper to one of several tracked pairs of shutter glasses to get a full three dimensional impression of the object. But as the stereo view certainly hindered others looking at the object, it could easily be switched off again by putting the glasses down during discussions. While we do not display 3D objects in our content linking system, we use the concept of having virtual pieces of paper that can be moved around using physical devices such as the puck. The Shared Design Space [8], a system consisting of four projectors for an interactive surface on a tabletop and a projector to create an interactive wall, is of interest as it not only use video cameras to track objects for interaction. Anoto Pens, digital pens that can track a pattern of tiny dots on the surface they are writing on, are used to control the system as well as to draw onto the virtual documents (images). As we aimed for a simple interface to view existing documents, we don't provide such a feature at the moment.

3 Setting

As our instrumented meeting room plays an important role within the AMI project, the meeting assistant had to be integrated into this environment. It also meant that the new system should require as little changes to the meeting room as possible. Back projection systems, such as Microsoft's Surface Computers became quite popular recently, as they allow to track fingers and objects touching the display surface. However, they obviously require transparent tabletops as screens, or they come as complete units which would require significant changes to the existing room. Furthermore, as they require sufficient room below the surface for the projection as well as the video tracking of touches, they would not allow meeting participants to put their feet below the table and sit comfortably. For those reasons we chose to build an on projection onto the existing tabletops, even if that does not allow to track fingers touching the surface directly. As the system has to accomodate several meeting participants, we decided to use several projection units, one for each user.

4 User Interface

In order to make the interaction with the system as direct as possible, we aimed to make the projected objects graspable. We decided against hand tracking as it is difficult to distinguish between gestures meant to manipulate documents and gesturing during discussions. This is especially true as the current setup does not allow to detect if the user's hands touch the desk top. Instead we track physical objects that serve as interaction devices using the AR-ToolKit+ tracking library.

4.1 Document Handling

In order to move projected documents around, a physical object (paper grabber) is associated with them. As long as this connection exists, the virtual document follows the grabber. The grabber objects consist of a piece of cardboard containing three markers, one of them elevated on a box (see figures 1 and 1(a)).

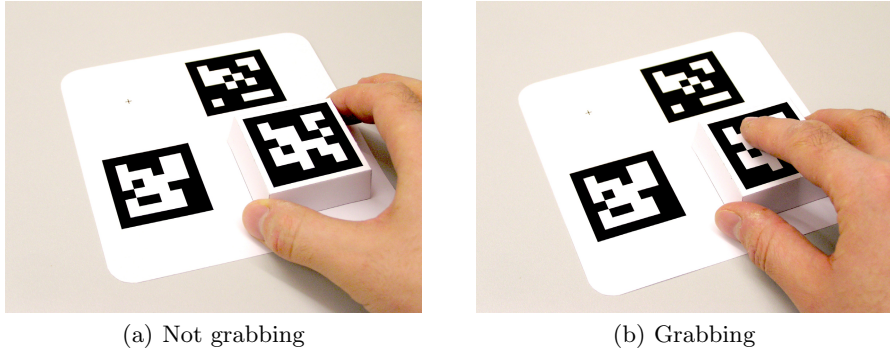


Fig. 1. The grabbing device and its switch marker

The elevated marker has the functionality of a switch. By blocking the visibility of this marker using a finger for example, one tells the system to grab the document below the device. If grabbed only on the sides, so that the marker on top is fully visible, the marker is disconnected and can be moved freely. Once the marker is placed on a virtual piece of paper, users can grab the document by holding the box like a mouse and thereby covering the top marker. Because the switch marker is on the box that user's grab, they do not have to think consciously about covering the marker or not. They just have to remember the two ways of holding the grabber device; On the sides to lift the grabber from the paper or with the hand on top of it, pressing it onto the paper they want to move. Once grabbed, the document stays connected to the grabber until it is released again, i.e. the top marker gets recognized again. This may be on the same or another user's projection system.

4.2 Keyboard Forwarding

Instead of providing virtual, projected keyboards as it is usually done with touch screen interfaces, we chose to use standard wireless keyboards. In order to allow keyboard based input, a Keyboard identified by the two markers attached to it (see figure 3(b)), can be placed on a displayed document. This allows to route keyboard events to the display applications that create the graphical representation of the documents. It replaces the physical connection (which keyboard is connected to which projection unit) with a virtual connection between keyboards and documents.

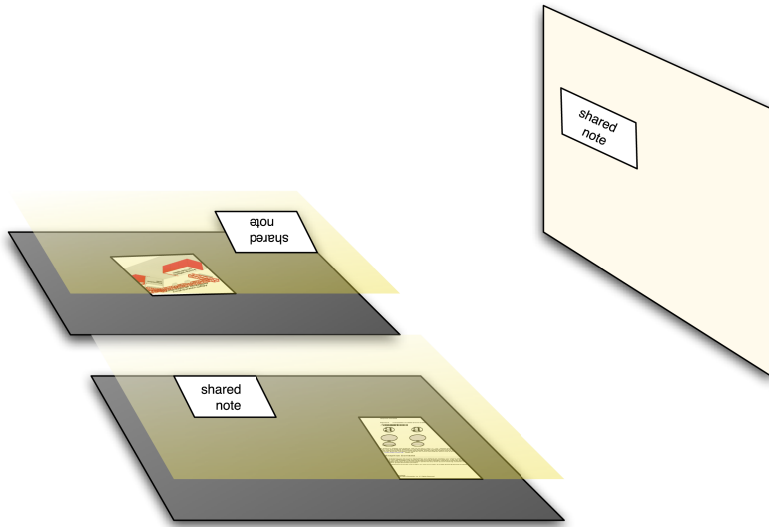


Fig. 2. The shared space as a semitransparent layer on top of the normal space. It can be (de)activated by covering a switch marker.

4.3 Sharing

In addition to augmenting the table, we wanted the system to incorporate the whiteboard as well. This way, participants are able to interact with content on the whiteboard directly from their place and move content between their space and the shared whiteboard space easily. While Hyperdragging as described in [9] would allow participants to do that in principle, it relies on a laptop with a conventional interaction device such as a touchpad. Using hyperdragging therefore would work against our goal to let the computer disappear.

We believe it is better to 'bring the shared screen to the participant' on the press of a button, or in our case when a marker is covered by the user. We therefore implemented a shared semitransparent layer (see figure 2) on top of the normal projection area which can be activated and deactivated by covering a marker placed on the projection area for that purpose. The presentation screen is the only exception here, as it does not have a private layer. It always displays the shared layer. Documents can be moved between the private and shared layers by grabbing them on one layer before switching to the other one. Once on the shared layer, all state changes such as position, orientation or which page of a multi page document is shown are forwarded immediately to other systems displaying the shared layer.

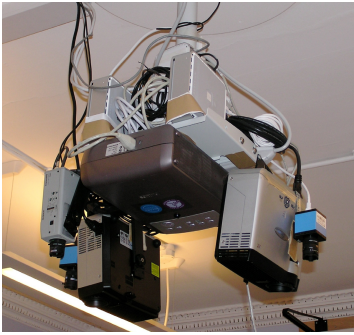
4.4 Auto-arrangement and Auto-iconizing

The Content Linking Device brings up new documents in regular intervals. In fact it often finds several documents to be displayed at the same time. In this situation

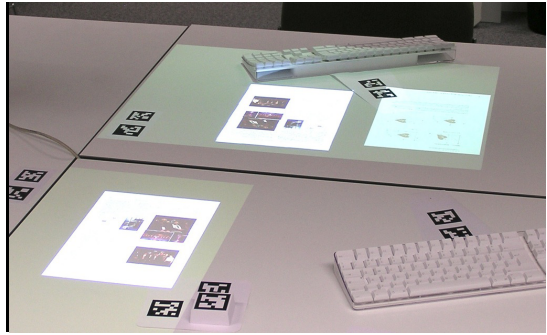
it is obviously not enough to make the documents appear at a fixed location (eg. the center of the tabletop). Of course the space on the table is not unlimited, so a method had to be developed to prioritize documents and remove less relevant ones gradually. We implemented a system to arrange and iconize documents automatically. It behaves as follows: New documents, are appended to an array of automatically arranged documents. If the number of elements in this array grows above a given limit (two documents in our case), the first one is removed and appended to the array of iconized documents. Additionally, a timer is set for each document added to the array of automatically arranged documents. Once the timer fires, the document gets iconized as well. This gives meeting participants the chance to have a glimpse at them and makes sure that documents that do not appear to be relevant to users are removed after 30 seconds. If the number of elements in the array of iconized documents grows above its limit (ten documents in our case), the first one, which has not been looked at for the longest, is removed and closed. Whenever documents are added to or removed from these arrays, all managed documents are sent a new goal position and scale factor according to the array they are in and their position within that array. The first auto arranged document is displayed on the left side. The second (and latest) one is positioned next to it in the center of the projection, leaving the right side for documents the user places there to read. Their scale factor is 1.0 so they are displayed in full size. The automatically iconized documents on the other hand are scaled down to 0.3 times their size and arranged along the front edge of the table with the oldest one being displayed on the left and newer ones being added to the right. When documents are sent to new positions or receive new scale factors, they don't change to these values immediately. Instead they animate towards these values over a given duration (1.5 seconds seemed best). This way it becomes obvious when the ranking changes and it is easy to follow what is going on. This is very useful when a document that is already open is deemed to be relevant by the query aggregator again, as one can see it move from its previous (possibly iconized) position to the position of the newest document (center). If the user places a paper grabber or keyboard connector on top of a virtual paper, it prevents the paper from being affected by the auto arrangement/iconizing system. If placed on an iconized paper, the paper is also scaled up to full size again. Now the user may move the document to where it can be read conveniently without interference of the auto arrangement system. Once the user removes the paper handling device and/or keyboard connector from the projected document, the system will take responsibility for the document again and iconize it after 30 seconds to clean up the tabletop. For the case that the document the query aggregator determined as relevant is already displayed as the latest document or controlled by the user, a visual ping has been implemented. If pinged, a document visually vibrates for a short period of time. It is animated by scaling it slightly up and down from its original size using a sine function. The amplitude of this vibration is scaled down to zero within 1.5 seconds to fade out the effect smoothly.

5 Our Setup

As mentioned before, we designed our system to consist of one projection unit per meeting participant. For our first prototype we decided to support two participants, as that would allow us to demonstrate the distributed nature of our system with a minimal budget. Each of the two projection units consisted of a projector (Optoma EP709), a camera (ImagingSource DFK 21BF04) and a computer (Mac Mini). Figure 3(a) shows the two projection units mounted around the projector for the shared screen. We used a MacBook Pro (not in the picture) to generate the content for this projector. A weakness of our first prototype,



(a) Setup of our first prototype



(b) Projection by our second prototype, devices

Fig. 3. Projection units and projected user interface

the resolution of its projectors (1024x768), became obvious quite soon. While this resolution was ok for documents containing little text in large fonts, such a meeting agenda or Power Point slides, it was not sufficient to display regular text documents as a virtual sheet of A4 paper. In order to address the resolution problem we replaced one of our Optoma EP709 projectors with an Optoma HD800X, a projector capable of projecting full HD video (1920x1080). In order to get an even better resolution from this projector, we used it in portrait mode, effectively augmenting only half of the users' table space. As documents are usually printed in portrait format, this enabled us to make better use of the projected pixels. While the EP709 projector would give us 768 pixels in height, the HD800X in portrait mode gives us 1920 pixels for the same height, an increase in resolution by a factor of 2.5 per dimension. Figure 3(b) shows the the projection areas for comparison, the larger area created by the lower resolution projector on the opposite side of the table and the smaller area originating from the HD projector on the front side.

6 Software Architecture

In order to keep the number of projection systems scalable, we divided our projection system into two parts. "Smart Projector", an application running on

all projection units and a central "Projection Manager". This display system is connected to the Content Linking Device via a third application, the "Hub Interface". These applications, as well as the other components of the Content Linking Device, may all run on different machines as they communicate via network. In our test setup however, we ran all these components on the laptop that rendered the contents of the shared screen. Only the two instances of Smart Projector augmenting the users' places, which had to perform video tracking as well, ran on separate machines, the two Mac minis.

7 Results

We developed a scalable projection system to be used in meeting environments. The way it is set up allows for easy installation in existing environments. After all, the camera, projector and computer could be integrated into a single unit mounted above the tables. We implemented software components that allow for easy management and coordination of the projection units as well as a user interface based on tracked interaction devices. We demonstrated that it is easily possible to move documents around on one, as well as between different projection units or between private and shared spaces. Figure 3(b) shows both projection units displaying a document on the shared space, while the unit on the opposite side shows a (slightly dimmer) local document as well. Furthermore, the system is able to connect to the central Hub of the AMI/AMIDA project. This way it can be used to display documents the content linking device deems relevant for the ongoing discussion. Additional functionality to manage the displayed content automatically was implemented to cope with the stream of new documents being introduced by the content linking device.

8 Future Work

In order to be able to evaluate the usefulness of the system, we will have to support more users with it, as the typical AMI/AMIDA meetings have four participants. While we believe that the higher resolution, which we achieved for one user in our second prototype, is necessary for every participant, we are not sure that augmenting half of a user's space is sufficient. In some situations a participant may use a laptop and other objects as well, so only half of one's space may be available anyway. In other situations however it may be important to have a glimpse at many documents at once, as they all could be important. In that case we would have to use two HD projectors per user. That however could lead to serious heat and noise issues in the meeting room.

Acknowledgements

This work is supported by the European IST Programme Project FP6-0033812 (AMIDA) as well as the Marie Curie Intra European Fellowship (IEF) FP7-221125 (NIPUI). This paper only reflects the authors' views and funding agencies are not liable for any use that may be made of the information contained herein.

References

1. Fapso, M., Schwarz, P., Szöke, I., Smrz, P., Schwarz, M., Cernocký, J., Karafiát, M., Burget, L.: Search engine for information retrieval from speech records. In: *Proceedings of the Third International Seminar on Computer Treatment of Slavic and East European Languages*, pp. 100–101 (2006)
2. Popescu-Belis, A., Boertjes, E., Kilgour, J., Poller, P., Castronovo, S., Wilson, T., Jaimes, A., Carletta, J.: The AMIDA automatic content linking device: Just-in-time document retrieval in meetings. In: Popescu-Belis, A., Stiefelwagen, R. (eds.) *MLMI 2008*. LNCS, vol. 5237, pp. 272–283. Springer, Heidelberg (2008)
3. Wellner, P.: The digitaldesk calculator: Tangible manipulation on a desk top display. In: *Proc. ACM SIGGRAPH Symposium on User Interface Software and Technology*, pp. 107–115 (1991)
4. Wellner, P.: Interacting with paper on the DigitalDesk. *Communications of the ACM* 36(7), 86–97 (1993)
5. Underkoffler, J., Ishii, H.: Illuminating light: An optical design tool with a luminous-tangible interface. In: *CHI*, pp. 542–549 (1998)
6. Underkoffler, J., Ullmer, B., Ishii, H.: Emancipated pixels: real-world graphics in the luminous room. In: Rockwood, A. (ed.) *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pp. 385–392. ACM Press/Addison-Wesley Publishing Co. (1999)
7. Ehnes, J., Knöpfle, C., Unbescheiden, M.: The pen and paper paradigm supporting multiple users on the virtual table. In: *Proceedings of the Virtual Reality 2001 Conference (VR 2001)*, p. 157. IEEE Computer Society, Los Alamitos (2001)
8. Haller, M., Brandl, P., Leithinger, D., Leitner, J., Seifried, T., Billinghamurst, M.: Shared design space: Sketching ideas using digital pens and a large augmented table-top setup. In: Pan, Z., Cheok, A.D., Haller, M., Lau, R.W.H., Saito, H., Liang, R. (eds.) *ICAT 2006*. LNCS, vol. 4282, pp. 185–196. Springer, Heidelberg (2006)
9. Rekimoto, J., Saitoh, M.: Augmented surfaces: A spatially continuous workspace for hybrid computing environments. In: *Proceedings of CHI 1999* (1999)